



Analysis Of Pedestrian Delay Models At Signalized And Unsignalized Zebra Crossings Along The Hyderabad Metro Rail Corridor

ANUKESH CHETLAPALLY¹, Sr. PROFESSOR KUMAR MOLUGARAM²

¹Ph. D (Full Time Scholar)

²Supervior Department of Civil Engineering
University College of Engineering
(Autonomous)

Osmania University, Hyderabad - 500 007

Abstract : Pedestrian delay is one of the concerning or related topics to pedestrian behavior at signalized intersections. When people try to cross the road, they may maintain the traffic signal and may not follow for some reasons. Those are there is a chance may be more crowded while crossing the road or maybe they are using mobile phones or maybe they are carrying more language with them. So, these all are having the chance of pedestrian delay. But the main problem is when serious accidents happen due to the delay of pedestrians, at that time, it is required to identify whose mistake it is and the reason behind that. Apart from that to reduce the crossing time or to save time, people sometimes do not follow the traffic signals also and this one more observed in teenagers. The majority of delays were decreased by those who started crossing during the flashing Don't Walk phase, even though only a minor portion of delays were reduced by beginning to cross when the signal was formally red. Compared to the number of people who would have arrived at the curb at the same time if arrivals were random, fewer pedestrians approached crosswalks during the flashing Don't Walk phases. Some pedestrians are expected to move quickly to enter during the flashing Don't Walk phase instead of waiting for the next Walk time. To provide a more accurate evaluation of pedestrian delay at signalized crossings, a delay equation was updated. Here the author created a model which can analyze pedestrian behavior and calculate the delay time based on the behavior of the pedestrian by using deep learning and open CV techniques. For that, the author collected the data through different live video footage of pedestrians from different sites and locations. Here the author focuses more on the school and bus station zones because these areas have more chances to delay due to the language carriage, college or university zones, and many more.

IndexTerms–Pedestrian, time, accidents, language.

I. INTRODUCTION

Pedestrian behavior is one of the biggest challenges and important factors nowadays because due to the increasing number of vehicles, there is an increasing chance of more accidents and due to this more life-taken accidents are going on. For that reason, every country is focusing on this particular issue more and trying to find advanced technology to control or analyze pedestrian behaviors. Here the author tries to create a tool or model by which it is easier to understand pedestrian behaviors and avoid or reduce accidents. After different research, it is obtained that pedestrian delay is the most common factor in signalized indications way. Because sometimes it is observed that due to various reasons, some pedestrians are walking very slowly and for that reason also most of the accidents are occurring day by day somewhere people are not waiting for a proper signal, and for that reason, accidents are occurring. But if it is unsignalized when the accidents occur then it's not impacted the drivers because it's a pedestrian mistake not the driver, but if it is signalized and

accidents are occurring then it is required to measure or identify the proper victim for whom the accidents are occurring. To solve these problems the author tries to create a model in this research study. Till now by different research, it is obtained that most of the research is going on in unsignalized intersections, and Feng *et al.*, (2007) proposed that they are creating a model which can observe pedestrian behavior as well as calculate the gap acceptance between the vehicle and the pedestrian. However, Tang *et al.*, (2020) suggested that Traffic engineers are working to reduce traffic and increase intersection productivity. The symmetric intersection (SI), which is practicable, affordable, efficient, and feasible, can boost the capacity of intersections. There hasn't been any research done on SI pedestrian crossing designs. The effectiveness and safety of various crossing patterns are evaluated, and three pedestrian crossing designs are developed in this research. Proposed delay models incorporate both through and diagonal pedestrian movements to increase efficiency (Tang *et al.*, 2020). Conflicts between exposures and the likelihood of automobile accidents are examined for safety reasons. The entire cost is then computed once the delay and probable accidents have been transformed into the appropriate monetary quantities.



Figure 1 Unsignalized pedestrian behavior

Source: Golakiya and Dhamaniya, (2021)

Golakiya and Dhamaniya, (2021) suggested a model that can analyze the urban midblock crosswalk based on the vehicle delay and for that, they used various deep learning and open cv techniques to reach the optimal output. They also say that there are appropriate pedestrian crossing facilities in urban areas of developing nations like India. The most vulnerable road users have received less consideration while designing traffic infrastructure, which has mostly focused on motorized vehicles. As a result, people are free to cross and move through regular motorized traffic lanes without any restrictions. Pedestrians cross the street at any undesignated, that is, unmarked by zebra crossings or pedestrian signals, a convenient spot close to their destination. The risk of pedestrians being hit by moving cars grew dramatically and these pedestrian crossing techniques disrupted regular traffic (Rad *et al.*, 2020). These crossing pedestrians cause traffic delays, which hurt the section's capacity and stream speed. Zebra crossings and signalized crosswalks are two examples of pedestrian crossing infrastructure that are not covered by Indian standards. So, different literature review it is observed it is a crucial part of nowadays that's why the author tries to create the model to analyze the pedestrian analysis in a signalized way. But here the authors more focused on the delay part of pedestrians and tried to find the proper reason for the delay.

II. METHODOLOGY

There are two major parts present for solving the problem. The first one is the physical model, here person goes and measures everything using tools (Babapouret *et al.*, 2022). It is a time taking process because here all the work is done by humans. And there is also some drawback like some time measurement should be wrong due to lack of proper communication. Here first they did some experiments and after that, they observed the entire data using some testing. But in the Mathematical model, the entire scenario will be changed because here they used some smart tools for collecting data as well as observing the data. After that simulate all the data for getting the result. But one thing is both models used some scientific tools during processing time. Here the author mentioned how to calculate the proposed works like distance, speed, delay, etc. For calculating the distance, the author considers one point and based on this point measure every pedestrian distance who crosses the road. Speed also calculates based on the same point, here the author measures the walking speed of all the pedestrians during crossing time. Another thing is to delay calculation means here the author tries to find out the actual reason why pedestrians get late during the crossing period. There are so many reasons are present which are affecting the speed of walking time. For example, if the pedestrian has handicraft, then they required more time because of their health issues. But at the same time, if

the pedestrian has more baggage or heavy weight luggage then also, they have late. Many times, people using mobile phones also have the same issues.

In the above part, the author already mentions three measures distance, speed, and delay. Now the author describes all the calculation processes one by one. Because distance and speed both techniques help for delay calculation. Elen and Avuçlu, (2021) suggested that for calculating distance there are many techniques available like Euclidean Distance, Minkowski Distance, Manhattan Distance, Cosine Distance, etc.

A variety of methods used various distance metrics. For the K-Nearest Neighbours technique, for instance, the Euclidean Distance method was utilized to calculate the distance metrics. Cui and Xia proposed that SVM used Mahalanobis distance to assess the classification problem. It suggests that before selecting the distance measurement, it must be vital to understand both the data and the algorithms that perform best with it. Therefore, depending on the algorithms, everyone can choose the distance processes.

Because of this, the author in this instance used Manhattan, Minkowski, and Euclidean distance to evaluate the success of the endeavor based on several algorithmic techniques. Each element of the methods indicated below is covered by the author.

2.1 Distance measurement

2.1.1 Euclidean Distance

The Euclidean distance in mathematics is the distance between two points. In other words, the length of the line segment connecting any two locations in Euclidean space is the Euclidean distance between them (Wang *et al.*, 2021). Even though Euclid did not refer to distances as numbers and the Pythagorean theorem was not connected with distance calculation until the 18th century, these names are derived from the ancient Greek mathematicians Pythagoras and Euclid.

The Formula of Euclidean the Distance is

$$= \frac{a_2 - a_1}{n} = \frac{b_2 - b_1}{n} \quad \text{where } a, b = \text{two points in Euclidean } n\text{-stage}$$

Euclidean vectors, starting from the basis of the stage (initial stage)

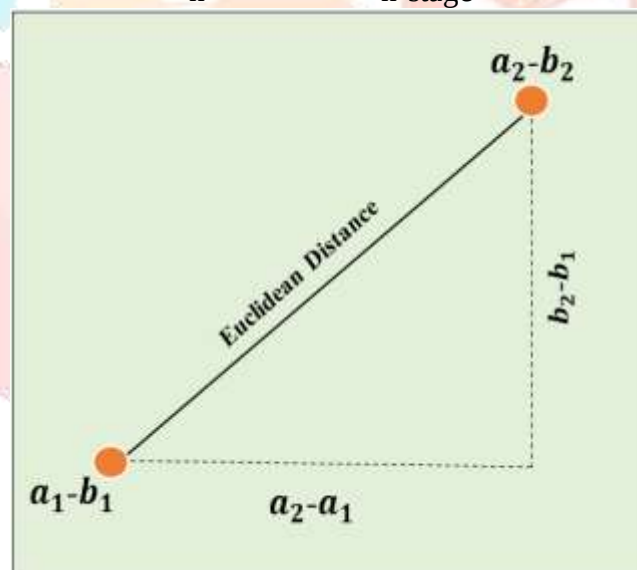


Figure 2 Calculation of Euclidean Distance

Source: Own Diagram

The main benefit of Euclidean distance is that it provides it easy and predictable to measure the distance between two places. Additionally, it gives single observations and calculates the distance between two places. For this reason, almost all of the studies used this method to calculate the distance. This approach does have certain disadvantages, though. The distance from two points it estimates could be skewed and dependent on other feature units because it is unable to scale in-variant data. Additionally, the dataset needs to be normalized before using this distance-measuring technique.

2.1.2 Manhattan Distance

The L1-norm of the Manhattan distance is $M(a, b)$ by (Veerashetty and Patil, 2021). The Manhattan Distance is the distance between two places that are separated from one another by right-angled axes. The Manhattan distance is the distance that a car would need to go in a city with square-blocked structures and right-angle intersections of straight streets like Manhattan. This also defines phrases like "city block" and "taxicab distance." The Manhattan distance measures the same distance as the Euclidean distance, but it averages the pairwise absolute difference rather than the squared difference between each variable. $M =$

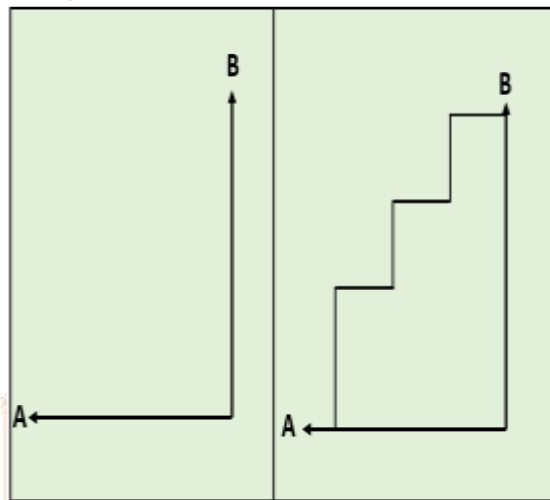


Figure 3 Calculation of Manhattan Distance

Source: Own Diagram

Manhattan Distance's primary benefit is how simple it is to compute the model. It always finished the model's resolution using a grid (Coudronet *al.*, 2020). In this strategy, any path is valid. This method also has several shortcomings, such as the fact that all of the roads are not singular and that there are different turns. The number of turns is always reduced while choosing a route. One of the most crucial elements is that the path is changed when arbitrary barriers are nearby.

2.1.3 Makowski Distance

The estimated distance between two points in N-dimensional space is known as the Minkowski distance. The Manhattan and Euclidean distances are both effectively generalized by this distance (Mailagaha and Luukka, 2022). In machine learning, it is commonly used, especially when attempting to find the best accurate classification or correlation of data. It establishes the spatial separation between two or more vectors. The Minkowski equation is used to construct the distance metric in machine learning, which is used to compare the sizes of two objects. It is also known as a "p-norm vector," which designates the order of the norm.

Minkowski Distance equation is, $M(A, B) =$

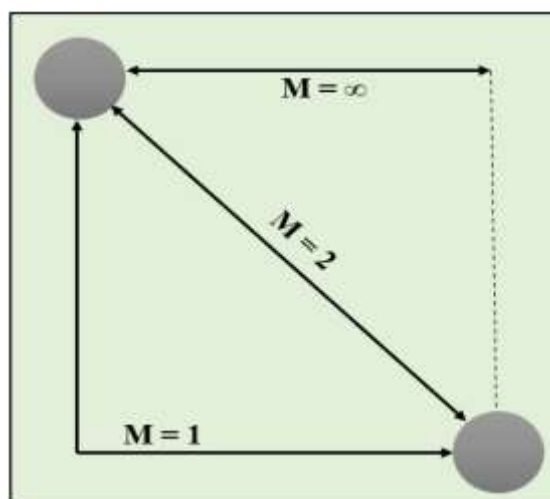


Figure 4 Calculation of Makowski Distance

Source: Own Diagram

The values of M are:

- $m = 1$ — Manhattan distance
- $m = 2$ — Euclidean distance
- $m = \infty$ — Chebyshev distance

Minkowski has the same restrictions as the distance measures that phrases like Manhattan, Euclidean, and Chebyshev distance represent, therefore understanding them is crucial. However, the inventor of the project solely employed the Manhattan and Euclidean distances to determine its distance.

2.2 Speed measurement

Here when a pedestrian entered the area of the zebra crossing then it generated a unique id for every pedestrian. So, when it entered the range, it stored the initial distance as well as recorded the distance coverage by the pedestrian. So, it records the distance as well as the time also. So, with the help of these two key values, the author calculated the speed of the pedestrian. With the help of this formula, the author calculated the speed of pedestrians,

$$\text{Speed} = \text{Distance}/\text{Time} \quad .$$

Where the speed formula is abbreviated as,

S = speed

d = distance traveled

t = time elapsed

2.3 Delay measurement

The author chooses a different area of collecting the data for this research work because of proper data collection and getting the appropriate data of carrying language in bus station zone, using mobile phones while crossing the zebra in educational zones, due to some medical issues and physical problem such as handicap peoples are also taking more time means a delay in hospital zones such data collection done by author by using different areas. Here the average speed of a pedestrian is 1.265 m/s so if a person has less speed while crossing the zebra it is considered as a delay. For that here the author used some of the speed calculation formulae to find the speed of the pedestrian shown in section A. Here the author tries to calculate the distance or length of the zebra crossing because the length of the zebra is required to calculate the speed of the pedestrian. After that, the author also calculated the time taken by the pedestrian to cross the road because if the author divided the length by the time taken then it is possible to calculate the speed and then comparing to the normal speed and pedestrian speed the author can decide whether a pedestrian is a delay or not

III. RESULT AND ANALYSIS

The numerous detections and the outcomes of the pedestrian crossing time reasons for the delay will be covered in this section. To produce the result, the analysis is divided into various tasks. The author will be predicting the age and Gender of pedestrians, time head and space head, speed, and distance for this proposed study. First, the number of pedestrians at a particular time was determined. Second, the distance is impacted by the measured time. Finally, the information was shown on the screen and the pedestrian was located.

3.1 Detection of Gender and Age

This technology can be applied when pedestrians are detected. The system accurately detects the pedestrian. An endless number of pedestrians could be efficiently recognized at once. The COCO dataset from internet pages was utilized for the prediction goals of this author. The results analysis was successfully finished. This system was installed to recognize abrupt pedestrian crossings, which happen when a few people cross in front of a moving vehicle without the driver noticing. The use of this gadget would serve to inform the driver of the pedestrian crossing. The notice "Road status: Safe Zone" will appear if the lane is safe for the zebra crossing. The person is carrying the luggage and crossing the zebra line. The male has nearly 25-32, the age of the pedestrian that person is in the safe zone, and carrying extra items of luggage can be delayed in crossing the zebra. In the second image, the person is describing the detection of Gender-based and also the age of the person, the person has 8-12 years male, and the pedestrian is in a safe zone when crossing the zebra line.



Figure 5 Detection of age and gender detection with luggage

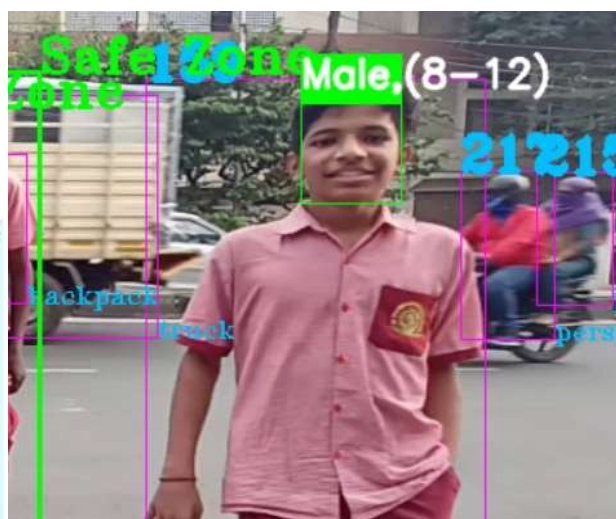


Figure 6 Detection of age and gender detection of schoolboy on zebra cross

3.2 Speed, Time head, and space head

From both figures, the school children are crossings the zebra line from one place to another place. In another image, the person is crossing the road, and beside the pedestrian crossing the vehicle of the driver is about 19 feet distance and has to reach 14 feet for their destination. The distance between the pedestrian and the vehicle is known as the distance, and the pedestrian to cross and reach the time taken period is known as time head. When the signalized and unsignalized on traffic, when a signalized pedestrian is in the safe zone, the unsignalized on traffic pedestrian will get suddenly may happen an accident, then reach their destination to be delayed.



Figure 7 Detection of speed at Zebra cross



Figure 8 Detection of time head, and distance head

3.3 On Zebra and Of Zebra Crossings

One of the main traffic laws is that all pedestrians must cross at the zebra, and they should do so at all times. There may be mishaps if the pedestrian does not cross the zebra. Because of this, the author predicts whether individuals crossing the zebra did it by the law. Here, the author employs the polyline technique to tally the number of pedestrians on and off the zebra so that the reader can comprehend the distance between automobiles and pedestrians and whether or not the pedestrian is in a safe zone. On the road, the pedestrian walking on the zebra line has 24 and the out of zebra has 41 pedestrians. When the pedestrian crossing on the zebra carrying luggage is in the safe zone, but to reach their destination, it is delayed.

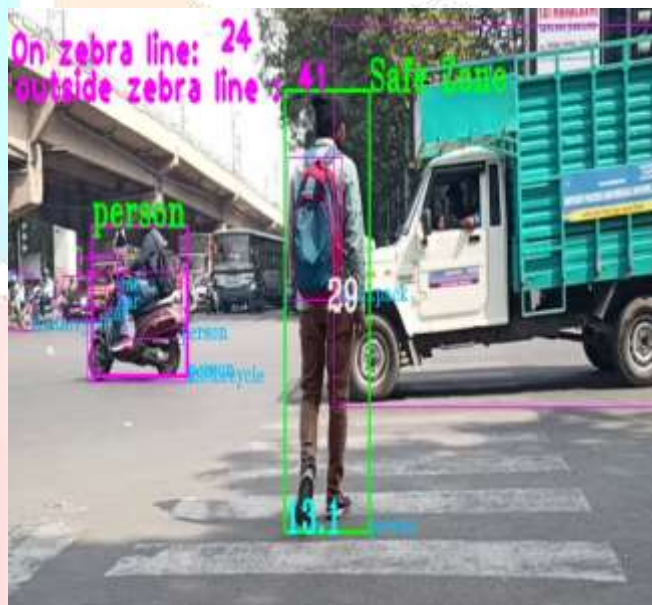


Figure 9 Detection of and on the zebra line crossing

3.4 Distance Between Pedestrians and the Endpoint of the Zebra

The author uses a few logical and mathematical questions to determine the distance between the pedestrian and the endpoint of the zebra crossing before using Opencv to display the research's findings. This section of the article predicts the distance between the pedestrian and the endpoint of the zebra crossing. The author measures the amount of time required for crossing in this section and establishes the goals based on the previous segment. The person or thing that was slowly moving across the zebra may have been carrying big bags and handbags for this reason. From the above figure, 13.1 feet for the zebra crossing to reach their destination. The pedestrian having the luggage and bags carried then pedestrians will be delayed.



Figure 10 Detection of the endpoint between the pedestrian reaching the zebra line

After collecting the data from the open cv the author going to display it in the table.

Table 1 Signalized and unsignalized area pedestrian count

Zone	NO	YES
Bus station	345	0
Industrial	178	613
University	204	629
Shopping mall	344	927

In Figure 11 below, which shows the pedestrian count, this table 1 contains a zone-by-zone breakdown of the pedestrian population.

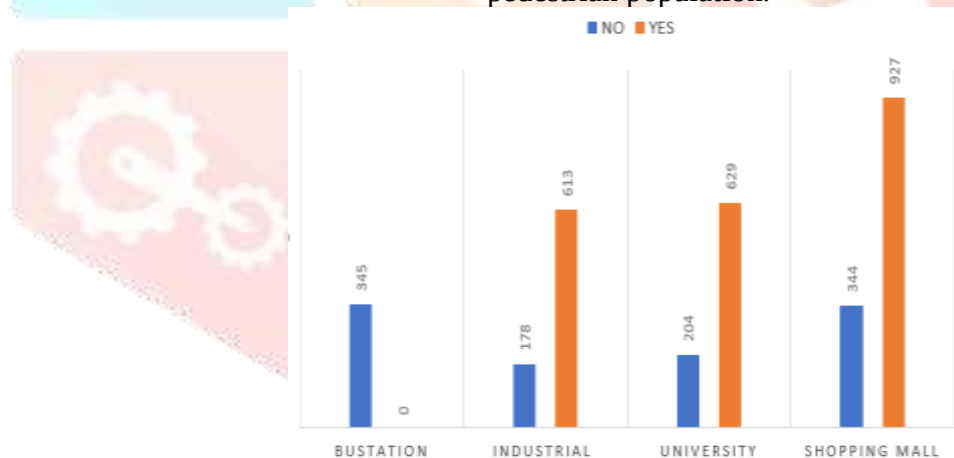


Figure 11 Pedestrian count in zone wise

Table 2 Causes of delay count

ZONE	femal e	mal e	offlin e zebra	onlin e zebr a	using mobil e	no mobil e	signalize d	un signalize d	languag e	no languag e
Bus station	170	175	171	174	184	161	345	0	173	172
Industria l	390	401	253	538	388	403	178	613	412	379
universit y	409	424	327	506	403	430	204	629	406	427
shoppin g mall	620	651	346	925	625	646	344	927	655	616

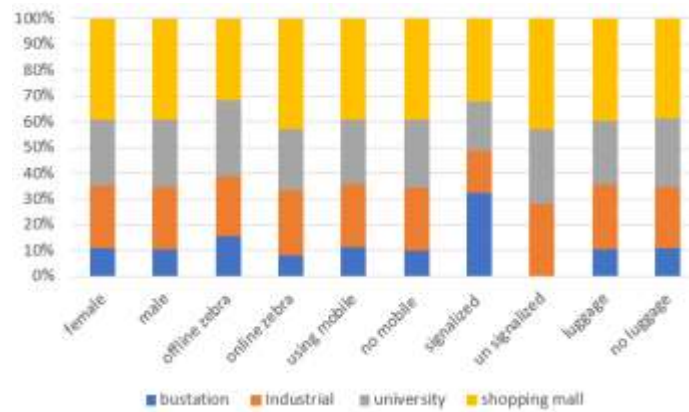


Figure 12 Pedestrian count in zone wise

The counting of walking on the road is shown in Fig. 12. The main reasons why pedestrian delays occur are been covered in the paragraph above, and this graph demonstrates how many individuals respect the regulations.

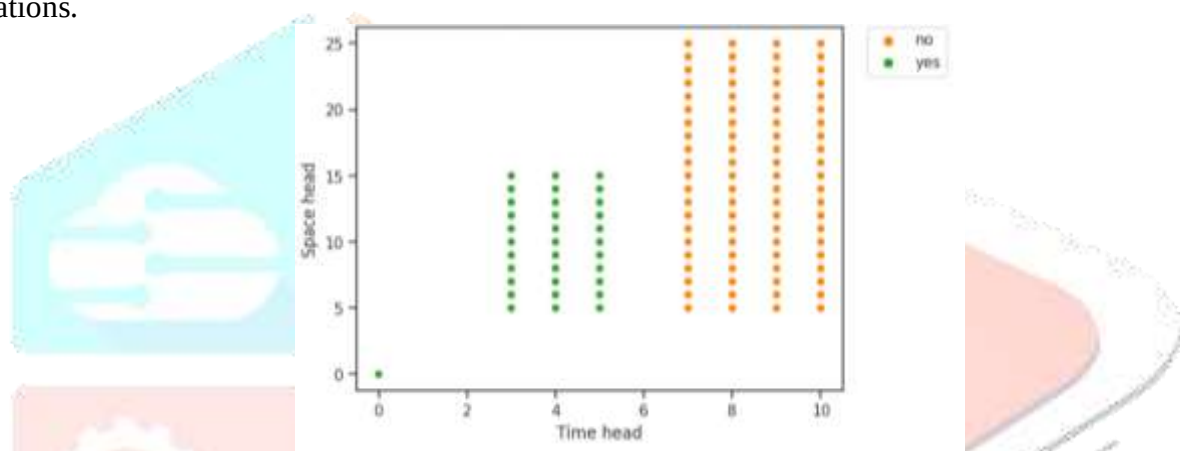


Figure 13 Relationship between time and space head.

Fig. 13 shows the delay of a pedestrian crossing a road. Green indicates a signalized area, whereas orange denotes a non-signalized area. Compared to un-signalized areas, signalized areas have no delay and low pedestrian time and space heads.

Table 3 AI model performance

Model	Accuracy
KNN	94%
Decision tree	99%
Random forest	99%
XG boost	99%

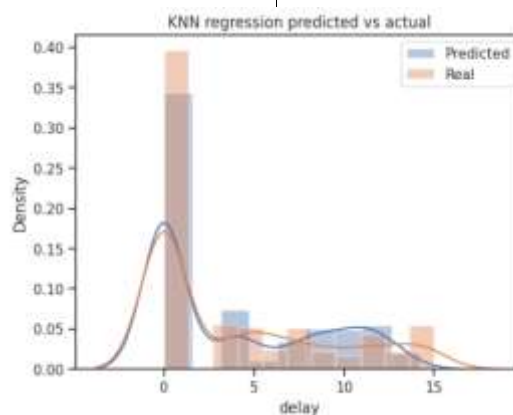


Figure 14 KNN model prediction

Figure 14 shows the prototype. The reader may comprehend that the blue colour reflects the forecast and the red colour the actual value in this KNN is how much it predicted section. Here, the reader may quickly see the difference between the prediction and real value. KNN is one of the lazy learner predictors; it predicts with 94% accuracy and is appropriate for anticipating pedestrian delays.

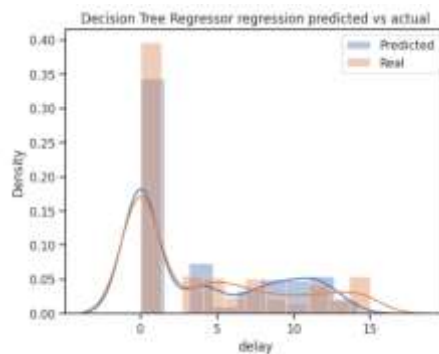


Figure 15 Decision tree model prediction

The discrepancy between the model's forecast and the actual is depicted in Figure 15. With a 99% accuracy rate, this model is slightly overfitting when compared to other models.

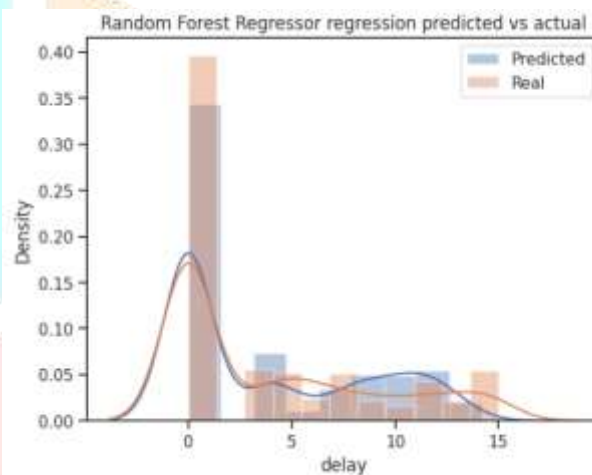


Figure 16 Random Forest model prediction

A random forest regressor that predicts the delay with 99% accuracy is shown in Figure 16. Additionally, the model is being overfitted when almost the same results are predicted.

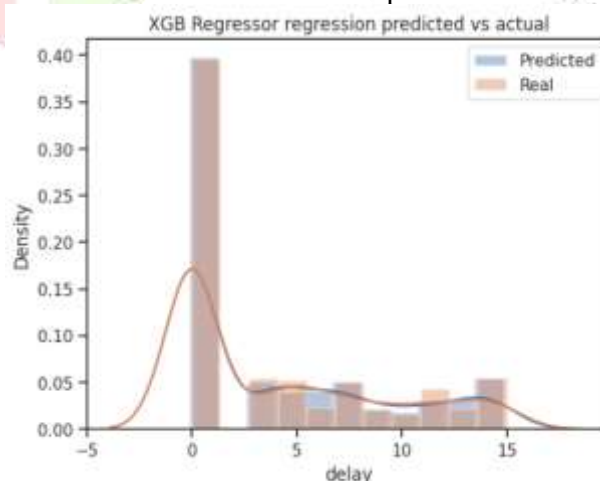


Figure 17 XG boost model prediction

The performance of the XG boost regressor, another ensemble technique, is superior to the random forest and the decision tree and predicts with 99% accuracy.

Table 4 KNN prediction result

Actual	Predicted
7.0	8.8
4.0	3.8
8.0	7.4
15.0	12.0
0.0	0.0

Figure 14 displays the small difference between the model's forecast and the actual data. The model was well-trained and outperformed other models in terms of preference based on the behaviour of the pedestrian.

Table 5 Decision tree prediction result

Actual	Predicted
7.0	7.0
4.0	4.0
8.0	8.0
15.0	15.0
0.0	0.0

Real-world and anticipated values given that both are comparable and the model is overfitting, this type of forecast is meaningless.

Table 6 Random Forest prediction result

Actual	Predicted
7.0	8.8
4.0	3.8
8.0	7.4
15.0	12.0
0.0	0.0

The random forest predicting table demonstrates that there is a small difference between the actual and model prediction values; this type of prediction is also good.

Table 7 XG Boost prediction result

Actual	Predicted
7.0	7.022
4.0	3.99
8.0	7.968
15.0	14.17
0.0	-0.0051

This kind of result is excellent for forecasting the delay because it is predicted by the XG boost with great accuracy. For future forecasting, a disparity of this size is preferable

The causes of the pedestrian delay are shown in Table 2, where the reader can see which pedestrians respect the rules of road traffic and which ones are challenging to cross the road. These key factors are mentioned here: gender, zebra crossing, mobile use or not, and whether or not they are carrying luggage

prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

IV. CONCLUSION

Finally, this study successfully identifies pedestrian crossings at predetermined distances. It might also be a response to a search for numerous individuals. There is a delay on the LCD panel after the systems recognize the pedestrian. Overall, this device is capable of detecting unexpected pedestrian crossings in full. It can be used by any vehicle that already has this technology installed. Apart from that it can be helpful for both vehicle drivers as well as pedestrians also because it can calculate the delay time as well as the distance between the pedestrian and the vehicle also. The author has detected the Age and gender, Time head, space head, and of and on a zebra crossing, and endpoint distance between the pedestrian and zebra cross for the pedestrian's prediction. The pedestrian is walking on the So, if a person uses this technique in the vehicle so they can have the proper information or proof and traffic control people also get secure and useful tools by using this.

REFERENCES

- [1] Feng, S.M. and Pei, Y.L., 2007. Research on the delay of a pedestrian crossing. Harbin GongyeDaxueXuebao, 39(4), pp.613-616.
- [2] Tang, L., Liu, Y., Li, J., Qi, R., Zheng, S., Chen, B. and Yang, H., 2020. Pedestrian crossing design and analysis for symmetric intersections: Efficiency and safety. Transportation research part A: policy and practice, 142, pp.187-206.
- [3] Golakiya, H.D. and Dhamaniya, A., 2021. Development of pedestrian crossing facility warrants for urban midblock crosswalks based on vehicular delay. Transportation in developing economies, 7(2), p.18.
- [4] Rad, S.R., de Almeida Correia, G.H. and Hagenzieker, M., 2020. Pedestrians' road crossing behaviour in front of automated vehicles: Results from a pedestrian simulation experiment using agent-based modelling. Transportation research part F: traffic psychology and behaviour, 69, pp.101-119.
- [5] BabapourChafi, M., Hultberg, A. and Bozic Yams, N., 2022. Post-pandemic office work: Perceived challenges and opportunities for a sustainable work environment. Sustainability, 14(1), p.294.
- [6] Elen, A. and Avuçlu, E., 2021. Standardized Variable Distances: A distance-based machine learning method. Applied Soft Computing, 98, p.106855.
- [7] Cui, L. and Xia, Y., 2022. Semi-supervised sparse least squares support vector machine based on Mahalanobis distance. Applied Intelligence, 52(12), pp.14294-14312.
- [8] Wang, S., Tian, Z., Dong, K. and Xie, Q., 2021. Inconsistency of neighborhood based on Voronoi tessellation and Euclidean distance. Journal of Alloys and Compounds, 854, p.156983.
- [9] Veerashetty, S. and Patil, N.B., 2021. Manhattan distance-based histogram of oriented gradients for content-based medical image retrieval. International Journal of Computers and Applications, 43(9), pp.924-930.
- [10] Coudron, I., Puttemans, S., Goedemé, T. and Vandewalle, P., 2020. Semantic extraction of permanent structures for the reconstruction of building interiors from point clouds. Sensors, 20(23), p.6916.
- [11] MailagahaKumbure, M. and Luukka, P., 2022. A generalized fuzzy k-nearest neighbor regression model based on Minkowski distance. Granular Computing, 7(3), pp.657-671.